Commander, Operational Test and Evaluation Force

OTD Checklists

Program: _____

Date: _____

POCs for OTD Checklists

COTF 01B—Test Development & Design

Mr. Dave Beyrodt.....01B (xt 3168)
CDR Tom Cecil.....01B1(xt 3388)

CORE TEAM FACILITORS

Mr. Kris Bauer..... (AWOL)
Mr. Paul Coslick ..... (xt 3373)
Ms. Felicia Farrow ..... (xt 3179)
Mr. Chuck Gillum..... (xt 3096)
Mr. Scott Cook..... (xt 3112)

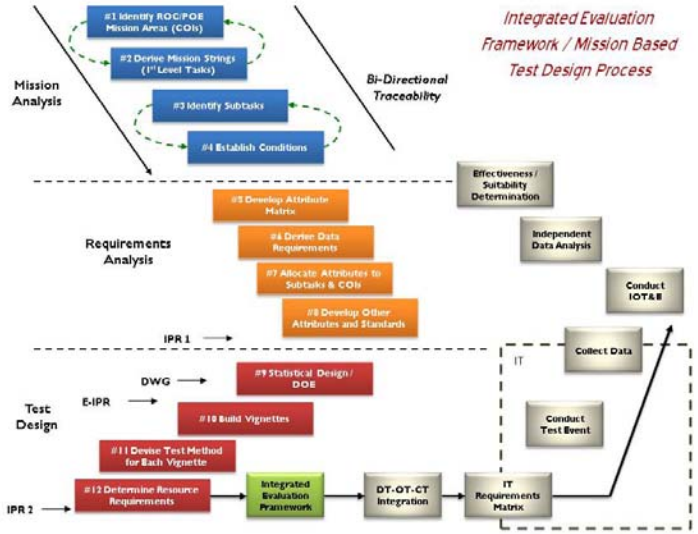
STATISTICAL DESIGN SUPPORT

Dr. Bob McIntyre ..... (xt 3132)
Mr. Josh Tribble ..... (xt 3280)

Version 1.7, February 2013

OTD IEF Checklist

Purpose: This is a step by step checklist leading an OTD up to IPR #1 and covers the first 8 steps of mission based test design (MBTD). It relates the initial MBTD steps to development of sections in the Integrated Framework.



At the conclusion of IPR #1, OTDs should move on to the Design of Experiment (DOE) checklist, conduct an E-IPR and then proceed to the IPR #2 checklist.

CAUTION: Using this checklist does not absolve the OTD of the responsibility for thought or the requirement to understand why they are performing these steps and how to apply the results. If you don't understand, ask.

Note: The OTD and the CTF should both sign the "Date Completed" blocks throughout this checklist. Use of the IEF database is required for all IEFs. Most of the tables in the IEF document are produced directly from this database.

OTD IEF Touchpoint 1 Checklist

Purpose: Building the reference library, defining System Under Test/System of Systems, initial mission analysis and COI selection

☐ 1. Gather and review the applicable reference documents for the system under test (SUT).

☐ a. ORD, CDD and/or CPD

NOTE: ***ORD = Operational Requirement Document (old school), CDD = Capability Design Document, CPD = Capability Production Document***

They are the overarching OPNAV and Joint Requirement Oversight Council (JROC) approved requirement document.

NOTE: ***OTDs need to be aware of the schedule and plan for updates to the CDD or CPD. The program office or OPNAV resource sponsor can provide this info. All OTDs must ensure COTF is aware of and is included in the review process. OTDs shall review their JCIDS requirement documents and submit recommended changes to OPNAV and the Program Office when they are routed for review.***

☐ b. Test & Evaluation Master Plan (TEMP)

NOTE: ***A program office document that requires COMOPTEVFOR concurrence. The program office T&E IPT lead manages all TEMP updates and is the primary source for updates and status of this document.***

☐ c. Concept of Operations (CONOPS)

NOTE: ***Usually provided by the user community and/or the resource sponsor (OPNAV). Brand new programs may not have one or may rely on an older one for a legacy system. Also consider any relevant TTPs.***

☐ d. Information Support Plan (ISP)

NOTE: A program office document that captures information technology requirements and interfaces in sufficient detail to enable testing and verification of those requirements. Some may contain useful mission task breakdowns for the system under test.

☐ e. Required Operational Capability/Projected Operational Environment (ROC/POE)

NOTE: Defines the primary mission areas and expected environment for the overarching platform. Not all programs will have a platform or system specific ROC/POE. These should align with the overarching USN ROC/POE – OPNAVINST C3501.2K (classified document).

☐ i. Platform specific ROC/POE

☐ ii. USN ROC/POE

NOTE: Unclassified excerpt of the 20 mission areas and first level operational capabilities is available on the Y: drive.

☐ f. Functional Requirement Documents (FRDs)

NOTE: Used to document requirements, typically at the sub-system level. May also be used to capture requirements for upgrades/modifications.

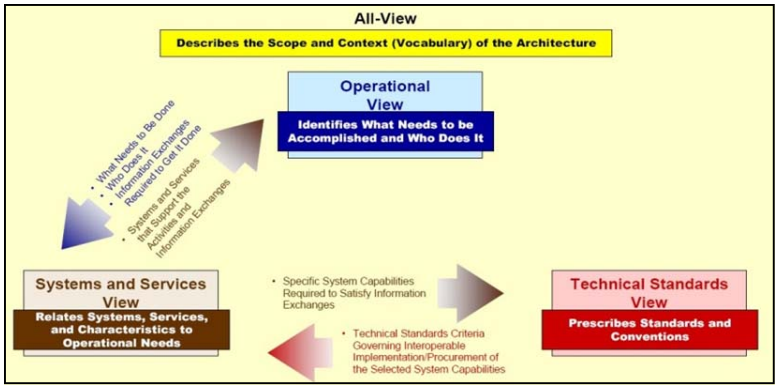
☐ g. DODAF Architectures

i. **OV-1 High Level Operational Concept Graphic**, typically found in TEMPs or the ORD/CDD/CPD

ii. **OV-5 Operational Activity Model**, depicts activities, relationships among activities, inputs and outputs.

iii. **OV-3/SV-6 Information Exchange Requirements (IERs)**, identifies the operational IERs between users and the SUT.

iv. **Other DODAF views as applicable (OV/SV/TVs)**



NOTE: DODAF architectures can be gathered through the program office or the resource sponsor (OPNAV).

☐ h. Review the CDD References and gather those deemed appropriate. References cited in a CDD/CPD are sources for specified requirements and should not be overlooked. (i.e. STAR, TTPs, SEP, TTVRs, etc).

☐ i. Security Classification Guide

NOTE: The Security Classification Guide is a required reference for every COTF test document.

Date Completed: _____

☐ 2. Define the System Under Test (SUT). This should be clearly defined in the CDD/CPD and/or applicable CONOPS.

a. Identify the final configuration of the SUT, to include major hardware & software components. If there are multiple phases of IOT&E/FOT&E with different configurations, explain that in the SUT definition.

b. Identify the end user (operator, maintainer, etc)

c. Identify the environment the system operates in

d. Describe the capabilities the SUT will provide, the capabilities gaps it will address and the desired effects of the system

☐ 3. Define the System of Systems (SOS). (This may be defined in requirement documents as Family of Systems).

a. Determine what other systems the SUT will interface and interact with that are outside the scope of test

b. Identify how the SUT impacts other systems and missions

NOTE: The SUT bounds the scope of test, but OTDs must be aware of the impact SUT deficiencies may have on the SOS. Understanding the relationship and definition of the SUT and the SOS will support categorizing deficiencies as Blue or Gold Sheets in the final report.

☐ 4. Draft the SUT and SOS Descriptions for inclusion in the IEF (Section 1). Use the applicable templates on the Y: drive.

NOTE: Readers should understand what the SUT is, who uses it, why they use it, how the SUT operates and supports the SOS.

Date Completed: _____

5. Determine the Effectiveness Critical Operational Issues (COIs)

☐ a. Review the OTD Manual, Chap 4 (Section 403)

☐ b. Review the Operational Capabilities for each Mission Area in the ROC/POE that could apply

☐ c. Identify the Mission Areas that capture the majority of operational capabilities affected by the SUT

NOTE: If the operational capabilities captured under multiple mission areas are similar and there is little difference in how the SUT is used (the tasks that operators perform are the same), then select the most stressing mission area as your COI(s).

CAUTION: If a review of the operational capabilities does not result in a mapping to specific ROC/POE mission areas, and a functional COI is considered approval must be received from 00 or 00D before proceeding. 01B CTFs can assist in this decision.

☐ d. Determine if an Information Assurance (IA) COI is required

NOTE: If the system is net-enabled (has connections and shares data with other systems over a network), an IA COI is required. 01C IA analysts can assist in this review.

☐ i. If an IA COI is required, use the standard COI language and IA references located on the Y: drive under IA Best Practices.

☐ ii. Review Standard Operating Procedure 10-01, OT of IA, for additional details

Date Completed: _____

☐ 6. Write the Effectiveness COI questions based on the mission areas selected.

a. Example: Will the (SUT) successfully accomplish (primary operational capability) in support of the (COI) mission?

Date Completed: _____

☐ 7. Determine Suitability COIs

a. Reliability, Maintainability, Availability & Logistics Supportability are normally standard COIs

b. Review optional suitability COIs identified in the OTD Manual

i. Use of optional suitability COIs requires approval from COTF 00 or 00D

ii. Items previously captured in optional COIs (i.e. Training, Documentation) should be evaluated under the appropriate effectiveness or suitability COIs.

☐ 8. Create Suitability COI questions.

Date Completed: _____

☐ 9. Draft the Effectiveness and Suitability COI sections of the IEF (first two sections of Chapter 2).

Date Completed: _____

☐ 10.Touchpoint #1

☐ a. Schedule a review with the 01B CTF, 01B A/B code, & division A/B code to approve COI selection (or 00/00D if a functional COI is proposed).

b. Be prepared to provide the relevant requirement document (CDD, CPD, etc) at this review for reference

c. OTD should provide:

i. Section 1 of the IEF

ii. Proposed COIs

iii. COI questions

Date Completed: _____

Attendees:

NOTE: If the approved COIs do not match those documented in an earlier version of the TEMP, a TEMP page change may need to be initiated.

OTD IEF Touchpoint 2 Checklist

Purpose: Developing the subtask hierarchy, defining conditions & beginning to build the IEF MBTD database.

Note: Use of the IEF MBTD database is MANDATORY. All subsequent steps should be captured in a program database.

☐ 1. Define 1st level tasks for each COI

☐ a. Review the OTD manual, section 403, which provides guidance on mission tasks.

☐ b. Review the default 1st level tasks provided for each COI. Identify the 1st level default tasks that are not affected or relevant to the SUT. These will be retained in your hierarchy, but grayed out.

☐ c. Identify additional 1st level tasks to consider adding to the COI mission threads as appropriate.

☐ d. Meet with the 01B CTF to create the initial IEF database for the specific project and review proposed task hierarchy.

2. Decompose 1st level tasks into subtasks

Date Completed: _____

☐ a. Provide enough detail so that all major components of mission accomplishment are accounted for. The OV-5 may be a useful resource for this step.

☐ b. Review the draft 2nd/3rd level subtasks commonly used by other programs, and select or modify as appropriate within the IEF dB.

NOTE: The intent of subtask decomposition is to define the mission of the SUT in terms of specific user tasks, which will then be used to plan and execute OT events (vignettes). Task decomposition (how far to go) should support assigning SUT measures/attributes as well as conditions (discussed below) to individual tasks. Tasks have to relate to the SUT—tasks that only relate to the SOS should be scrubbed.

CAUTION: Do not state obvious operator functions, but focus on meaningful events. Avoid decomposing tasks beyond these meaningful events. Recommend proceeding beyond a 3rd level only by exception. Do not include details of the operating environment that should instead be captured as conditions.

☐ 3. Use IA COI Task breakdown, if appropriate

☐ a. If the SUT has an IA COI, use the default IA task hierarchy provided by 01C (detect, protect, react, restore). IA task guidance is available on the Y: drive.

☐ 4. Develop Suitability task hierarchy

☐ a. Review common task hierarchies used by other programs for suitability COIs, available in the IEF dB.

b. Reliability and Availability COIs generally do not have a task breakdown associated with them.

c. Maintainability COIs may have task breakdowns that are unique to the SUT and required maintenance actions.

☐ d. Expand Suitability task hierarchy as required

Date Completed: _____

☐ 5. Review completed task hierarchy

☐ a. The task decomposition should allow OTDs to logically design a sequence of events (an end to end mission event or a subset of a mission that will be repeated numerous times) in a vignette, that a tester will execute as a test event.

☐ b. Verify the final task hierarchy logically captures the major tasks the SUT or operators perform as part of the applicable mission area (COI).

Date Completed: _____

☐ 6. Establish conditions

NOTE: Conditions are things that may affect performance of the SUT or influence operator actions when using the SUT (e.g. sea state, clutter, network load, presence of jamming, target type). They are broken into 4 categories: physical environment, military environment, civil environment (rarely used), & custom-created by the OTD.

☐ a. Review the requirement documents and the 01B database of conditions used by previous programs as a starting point. The CDD may identify environmental conditions that define the SUT's operating envelope. The 01B database contains the standard conditions defined in the UJTL & UNTL task lists, and custom conditions created by OTDs for their specific SUT.

☐ b. Select the conditions that will impact SUT performance or operator actions from each of the four categories.

☐ c. Identify and define custom conditions as required.

☐ d. In the SUT IEF dB and with initial CTF assistance, associate each condition with the appropriate lowest level task in the task hierarchy.

NOTE: The resulting linkage should identify what things can influence the operator's actions and/or performance of the SUT. Having this traceability supports building a vignette, and subsequent planning associated with design of experiments (DOE).

CAUTION: Do not attempt to document every conceivable condition that could affect the SUT (e.g. sun spots). Using the OTD's subject matter expertise, identify those most likely to impact performance or those of most interest to the testers.

Date Completed: _____

7. Begin thinking about whether the conditions are controllable or recordable and how the OTD might control them.

NOTE: The identification of controllable vs. recordable conditions will impact follow-on design of experiments, test resource requirements and data collection requirements. Recordable conditions are items that can't be specifically controlled (e.g. sea state) but are critical to capture for post-test analysis purposes. Controlled conditions are items that can be controlled/adjusted as desired by the OTD during an operational test event (e.g. presence of jamming). These controllable conditions will be used to identify different conditional variations required to test the SUT in. (e.g. day/high alt/EO mode vs night/low alt/IR, etc)

NOTE: Identification of conditions as controllable or recordable in the IEF does not prevent OTDs from modifying or updating their test design at a later date. For example, data collected and analyzed during IT events may drive OTDs to a different conditions list which will be documented in the appropriate test plan.

☐ 8. Establish descriptors (levels) for each condition

☐ a. Review the 01B database of conditions and their associated descriptors used by other programs.

☐ b. Verify the default descriptors provided are appropriate for the SUT .

☐ c. Modify the descriptors as required. Descriptors (levels of the conditions) can be modified by OTDs

Condition/Factor	Descriptor/Level
Altitude	Low (0-10K)
	Med (10-20K)
	High (20K+)
Target RCS	Small (Drone) (X ft^2)
	Large (Acft) (Y ft^2)

☐ d. Pay particular attention to the definition of each descriptor/level. They should be operationally relevant and clearly bounded for purposes of data collection and post test analysis.

NOTE: Descriptors should be identified in quantifiable terms vice simply "easy/medium/hard". Testers and reviewers need to know how each of those qualitative terms are defined. Also note that breaking conditions up into descriptors/levels (which is done to support subsequent DOE) does not imply analysis across the whole spectrum will not be done, if applicable.

Date Completed: _____

☐ 9. When conditions have been selected and descriptors defined, generate the conditions directory from the IEF dB. (see IEF template)

Date Completed: _____

☐ 10. Schedule Touchpoint 2 with the 01B CTF, and A/B codes from 01B and the division to review the complete task breakdown, selected conditions, and their associated descriptors. OTD should provide the following from the IEF dB:

a. Table B-1 (Conditions directory)

b. Table B-4, with the conditions linked to tasks (measures will be blank)

Date Completed: _____

OTD IPR #1 Checklist

Purpose: Building the attribute matrix, linking attributes to the task hierarchy and conducting IPR #1.

☐ 1. Review the overarching SUT requirement document (CDD/CPD/FRD, etc).

☐ 2. Review the attribute matrix in the IEF template.

> ***NOTE: For purposes of this checklist, the overarching requirement document will be referred to as the CDD, but may be some other sponsor-approved requirement doc.***

Date Completed: _____

☐ 3. Identify attributes/measures documented in the CDD

a. Attributes/measures are all statements of required capabilities (commonly referred to as the "will/shalls")

> ***NOTE: Attributes are capabilities the SUT is supposed to provide. Measures are the specific metrics used to assess those attributes. It generally is a one-to-one correlation, but a single attribute may be decomposed into multiple measures as appropriate. See the below example.***

A29	Fuel Endurance (KPP)	ORD	TAB 5-1	M85	Flank speed/ endurance	28 kt for 6 hr or 5% of a 120 -hr period	Specified
		ORD	TAB 5	M86	Transit speed/ endurance	18 kt for 18 hr or 15% of a 120-hr period	

☐ b. Extract all attributes from the CDD, with proper references to the source document (paragraph, table, etc), and annotation as a MOE, MOS or System of Systems (SOS) attribute. KPPs should be noted as such in the IEF dB. Criteria (i.e. thresholds) used to assess each attribute/ measure also needs to be captured.

> ***NOTE: MOEs and MOSs apply to the SUT. A SOS attribute is any capability or issue not previously captured that is needed for SUT mission accomplishment from a system-of-systems perspective. e.g. when the SUT is a fire control radar or weapon system that relies on the accuracy of a radar track passed to it by a system outside the SUT. While the SUT may perform effectively on its own, when the accuracy of the track provided to it is taken into account, the overall SOS may not perform effectively. Determining what metrics should be categorized as SOS attributes depends on a well defined SUT & SOS.***

Date Completed: _____

☐ 4. Review the other applicable reference documents. Extract derived attributes from those alternate sources, being sure to capture the specific source

> ***NOTE: The relevant capabilities document may not be the sole source of attributes/measures. Other sources of operationally relevant attributes may be applicable CONOPS, system specifications, platform level ROC/POE, previous test plans and reports of similar systems etc.***

Date Completed: _____

NOTE: Not every single "will" and "shall" statement in the requirement document is worthy of including in the attribute matrix as a measure. OTDs should scrub the resulting matrix for duplicate or unnecessary measures.

☐ 5. For each attribute/measure, define the type of attribute it is (only three options exist)

a. Specified = Clearly documented in the CDD. Either a KPP, MOE, MOS or stated as an important "will" or "shall" statement

b. Derived = Not explicitly stated in the CDD but come from some other source document and are useful in assessing the SUT capability.

NOTE: Derived attributes are documented or derived from other source documents (i.e. CONOPS, TTPs, SUT specifications, military standards, OPNAV instructions).

c. Other = attributes/measures that will be used to evaluate the effectiveness and suitability of the SUT, but do not have a source document or reference.

Date Completed: _____

☐ 6. Populate the IEF dB with the SUT attributes

a. Be sure to include the source, measure, criteria and type of attribute for each one

Date Completed: _____

☐ 7. Identify Information Assurance attributes

a. Specific common IA attributes and measures applicable to all systems are provided in the 01C Best Practices folder

b. Develop additional attributes and measures (with the corresponding source, criteria and type) based on the IT used to complete the attribute table.

☐ c. Review the IA related attributes with the 01C IA analysts.

Date Completed: _____

☐ 8. Define the data requirements for each measure. This can be done as attributes/measures are entered into the IEF dB.

NOTE: Data requirements consist of three primary items: (1) Element (i.e. temperature, position), (2) Unit of measure (e.g. °F, lat/long in degrees, min, sec), & (3) Data source (i.e. the specific mission computer data log). A data requirement should be created for each source.

For qualitative data requirements, the same items are required. Examples of standard qualitative data requirements are in the IEF template. The unit of measure may be qualitative comments or a Likert scale. Data source should point to a survey sheet or log.

Date Completed: _____

☐ 9. From this master attribute matrix, identify the orphan attributes and those that are DT only (data won't be collected on by OT)

NOTE: Orphan attributes are attributes that OT will not collect data for or look at as part of our assessment of effectiveness and suitability.

NOTE: DT only attributes are those that remain relevant to OT, but rely on data collected during DT for verification. Our operational assessment of those attributes will solely rely on data collected or reported by DT. These attributes are maintained in the attribute matrix but denoted with a "(DT Only)" after the measure description.

Date Completed: _____

☐ 10. Scrub the resulting attribute matrix and the criteria used to evaluate each measure for clarity

a. The criteria in the documentation may be ambiguous, contradictory between multiple measures, or simply need clarification

☐ b. Request clarification from the resource sponsor as necessary.

Date Completed: _____

☐ 11. Print the OT Attribute Matrix from the IEF database

Date Completed: _____

☐ 12. In the IEF dB, link the measures to the appropriate subtasks in the previously generated task hierarchy

a. Measures should be linked to the lowest subtask applicable.

NOTE: OTDs should be approaching this from the question "What measures do I need to evaluate the ability of the SUT to perform this task?" and then select measures that answer that question from the attribute matrix. If the measure does not help answer that question, don't link them.

☐ b. After the initial linkage is made, review the resulting orphan attribute table which shows measures that aren't linked to tasks. If there are measures in the orphan attribute table that are relevant to OT, then go back and link them to the appropriate task.

Date Completed: _____

☐ 13. Identify the critical tasks.

a. Review the COIs, first level tasks, lower level tasks and subtasks to identify where most critical tasks are captured and designate them as critical in the IEF database.

NOTE: Critical tasks are selected from the previously created task hierarchy and are essential to mission accomplishment. If the operator or SUT is not able to successfully accomplish the critical task, this could potentially result in finding a COI to be not effective or suitable.

Date Completed: _____

☐ 14. Identify critical measures.

a. The measures linked to the critical tasks are candidates for designating as critical measures and should support the test objective identified above.

NOTE: Candidates should include KPPs, if the KPPs are operationally relevant, significant MOEs, MOSs, and other measures deemed important to the OT, etc.

☐ b. Review the list of measures that are linked to the critical tasks identified above and create the critical tasks to critical measure table (see table 2.1 in IEF template). Compare against the KPPs, MOEs, & MOSs specifically identified in the relevant requirement document, previous TEMP, etc. and verify that all critical measures are accounted for.

NOTE: Requirement documents are not perfect. There may be KPPs, MOEs or MOSs that are not of critical value in determining effectiveness and suitability of the SUT.

Date Completed: _____

☐ 15. Print out and review the resulting Attribute and Traceability Matrix from the IEF dB.

☐ a. The Traceability Matrix should show the task breakdown for each COI, with measures and conditions mapped to the appropriate subtask. Review for consistency and be prepared to defend the linkages.

☐ b. Verify that every subtask that is present in the task hierarchy has measures associated with it.

Date Completed: _____

☐ 16. Share the resulting tables with DOT&E and the program office T&E IPT lead if not done already

☐ a. Document their comments via email

Date Completed: _____

Date Comments Received: _____

☐ 17. Review IEF products and prepare for IPR #1

☐ a. Review IPR #1 brief template on the Y: drive

☐ b. Verify the following products are complete and ready for review:

i. COIs
ii. Task Hierarchy
iii. Conditions Directory
iv. Attribute Matrix
v. Orphaned Attribute Table
vi. Attribute Traceability Matrix (with linked measures/conditions)
vii. Critical tasks to critical measures table (2.1)
viii. Measure-to-data requirement table (available as a custom export from the IEF database)

☐ c. Build the IPR #1 brief

- [ ] d. Schedule IPR # 1 with Division A/B code, 01B A/B code, CTF, OTC, OTD, and contractor support personnel
    - i. DOT&E should be invited to participate

- [ ] 18. Conduct IPR #1

Date Completed: _____

- [ ] 19. OTD documents action items and shares with attendees

Date Completed: _____

- [ ] 20. Close Action Items

Date Completed: _____

21. Proceed to the Design of Experiments Checklist

OTD Design of Experiment (DOE) Checklist

(Step 9 of the MBTD/IEF Process: Statistical Design/DOE)

Purpose: Identifying Critical Tasks and Attributes, selecting response variables (RV), conditions/factors, developing the statistical design needed to generate a run matrix and determine the appropriate sample size to achieve a satisfactory statistical power associated with the RV or particular factors.

NOTE: Statistical Design of Experiments varies from one program to the next. One size does not fit all. OTDs need to work closely with their divisional analysts and 01B to arrive at a statistical design that is defendable and useful. The following checklist touches on the basics – each program will be unique. Communication amongst all parties, including with DOT&E and IDA reps is critical.

CAUTION: Using this checklist does not absolve the OTD of the responsibility for thought or the requirement to understand why they are performing these steps and how to apply the results. If you don't understand, ask.

- [ ] 1. Identify statistical test objectives. The test objectives should focus on assessing performance of the critical tasks.
    - a. Common objectives include:
        - i. Verifying performance is above a threshold across all conditions
        - ii. Verifying performance is above a threshold in a specific subset of conditions
        - iii. Verifying new system is as good as a legacy system

iv. Characterizing performance across the whole operational envelope (developing a design that supports ANOVA, if applicable).

☐ 2. Draft the "Critical Tasks & Measures" paragraph and table for each COI (see IEF template).

NOTE: This paragraph should identify the overall approach for evaluating each COI by pointing the reader to the critical tasks and measures used to evaluate that COI. The identification of critical tasks and measures does not imply that other measures mapped to that COI in the traceability matrix won't be looked at;, just that the critical one's carry more weight in the assessment.

Date Completed: _____

☐ 3. Review the list of critical measures and identify those that clearly should be evaluated via demonstrations. Analysis by demonstration usually applies to measures that will be evaluated qualitatively, under multiple conditions (multi-variate), or under recordable and/or constant conditions. Justification for demonstrations should be identified and defendable.

NOTE: There are several types of demonstrations: (1) Some measures deemed critical can be evaluated quickly via demonstrations and do not need multiple data points to evaluate (i.e. the ability to load the SUT onto a C-130 may be critical and can be verified by demo'ing it once). (2) Alternatively, demonstrating it several times using different support equipment would be a multi-variate demonstration. (3) Some demonstrations require multiple measurements on the critical measure or response variable. These measurements may be summarized by a summary statistic (e.g. mean or median) and an associated confidence inter-

☐ 4. Identify potential response variables from the list of remaining critical measures. Identify as many as required to cover the mission.

NOTE: Response variables are critical measures that are ideally quantitative in nature and must be analyzed with statistical methods to support conclusions in the report. They are used in planning to ensure a minimum adequate sample size (number of runs and/or data points) and in the pre-planning to analyze data and determine if that measure's threshold was met and/or how that measure's result changes due to controlled conditions (factors).

a. Response variables should be :

i. Testable (i.e. practical, able to collect data on)

ii. Documented – traceable to a requirement document (specified or derived)

iii. Valid (i.e., represent an essential aspect of SUT performance)

iv. Reliable (i.e., relatively free of random noise)

v. Meaningful – a direct measure of the mission performance we are interested in

NOTE: Ideally, response variables are explicitly identified and thresholded in a requirement document. In some cases, OTD's may derive response variables to better capture the SUT overall performance .

vi. Quantifiable (i.e. either a numerical performance measure or some qualitative characteristic for which a numerical scale can be directly developed)

vi. Preferably continuous vice discrete, focused on overall mission performance or key elements of a mission task breakdown (one or multiple critical tasks)

viii. Discriminating – should distinguish levels of effectiveness.

NOTE: Continuous variables can be plotted along a range of values on a numerical scale (e.g. time, range, speed). These are usually normally distributed, meaning the frequency of occurrence of values follows the familiar bell-curve which allow for the use of a variety of statistical analysis techniques.

Binomial variables are discrete yes/no, kill/no-kill, probabilities/proportions, etc. and do not provide operators with as much insight into the performance of a SUT in the intended environment. There are also other types of discrete variables (e.g. count data such as number of false alarms which usually follows the Poisson distribution), etc.

CAUTION: Using binomial/discrete response variables should be avoided, in favor of continuous variables if at all possible. While binomial variables can provide just as much info regarding system performance, they require significantly more data.

☐ b. For each response variable, determine if historical data from previous evaluations exist. This includes both previous operational testing and DT data. If available, locate and review with divisional and 01B analysts.

NOTE: Historical data are extremely useful in the subsequent DOE planning steps. They can provide a baseline for performance of a legacy system, validate assumptions in the numerical behavior of the measure (type of distribution, etc.), and serve as the basis for estimating the standard deviation of and effect size of expected test data, or screening of factors impacting the RV.

Date Completed: _____

☐ 5. Identify the conditions that are associated with the selected critical tasks and measures and the selected response variables. These should already be linked to the relevant tasks/subtasks and can be found by reviewing the previously built Attribute Traceability Matrix.

NOTE: This also serves as a sanity check of the conditions associated with the task. Previously, conditions were linked to tasks, and then measures were linked to tasks. Those conditions should also logically affect the measures if this was done properly. When thinking about conditions/factors that affect the critical measures, if there are conditions that affect those measures but weren't previously identified or linked to the parent task, then add them to the conditions directory and update the linkage. Do not overlook conditions that can't be controlled (recordable) but are important to collect data on to understand and analyze system performance.

Date Completed: _____

☐ 6. Prioritize the conditions associated with each response variable by the anticipated impact they will have on SUT or operator performance.

☐ 7. Prioritize the levels of each condition (which were previously identified as part of step 8 of the Touchpoint 2 checklist) as they apply to each critical measure.

☐ a. Estimate the effect that different levels of conditions have on the performance of the SUT as the condition changes between descriptors/ levels (i.e., significant/moderate/low effect).

☐ b. Estimate the likelihood of encountering the different levels in an operational environment. (i.e., all levels are equally encountered, some are seen more than others, etc).

NOTE: This step focuses the test design on the most operationally relevant environment/scenarios. Include outside organizations in this process. The goal is to ensure that the test design includes and focuses on the conditions that are most operationally relevant. This should be done for both controlled and recorded conditions.

c. Use the below table as a guide in assessing the levels of each condition.

		Likelihood of Encountering level During Operations		
		Multiple levels occur at balanced frequencies (e.g., 1/3, 1/3, 1/3)	Some levels are balanced, others are infrequent (e.g., 5/10, 4/10, 1/10)	One level dominates (e.g., 4/5, 1/10, 1/10)
Effect of Changing Level on Performance		**Balanced**	**Mixed**	**Dominant**
Significant Effect on Performance	**High**	Vary all	Vary balanced levels. Demonstrate infrequent levels	Fix dominant level. Demonstrate others.
Moderate Effect on Performance	**Medium**	Vary all	Vary balanced levels. Demonstrate others	Fix dominant level. Demonstrate others.
Low Effect on Performance	**Low**	Fix levels or record level used	Fix levels or record level used	Fix dominant level

d. The result of this exercise is the identification of factors (levels of conditions) that have an important effect on the performance of the SUT and are likely to be encountered by the operator. They will be used to design a test with statistical power and confidence. The levels of descriptors that have a low effect or are encountered infrequently may only be demonstrated.

e. When a single level dominates, testing may focus on the dominant level, with demonstrations for the other levels, if appropriate. Efforts should be made to define factors as continuous vice categorical.

Date Completed: _____

☐ 8. Meet and discuss proposed response variables, prioritized conditions and selected factors with divisional analyst and 01B CTF.

☐ a. Define the objective of the test

NOTE: Examples include (1) demonstrating performance is above a threshold across all conditions, (2) demonstrate performance is above a threshold in a specific subset of conditions, (3) demonstrate new system is as good as a legacy system, or (4) characterize performance across the whole operational envelope.

☐ b. Identify response variables and their associated thresholds

☐ c. Prioritize conditions with selected factors/levels using the above matrix

☐ d. Develop resulting matrix of all variations used for designing a statistical test. See below example:

Conditions		
Altitude	Airspeed	Target
0-5K	1-100	RCS A
5-10K	100-200	
10-15K	200-300	RCSB
15K+	300+	

NOTE: In this example, the OTD may have determined that low altitude, low airspeed, and airspeeds above 300kts will have a low impact on performance or are not likely to be encountered by the system. They may be considered for demonstration runs, but won't be included in the statistical DOE. The remaining levels will be used as factors in the design to meet statistical power.

☐ e. Identify disallowed combinations by reviewing the list of controlled conditions for combinations that are not testable or realistic (i.e. arctic terrain & hot temperatures, etc).

NOTE: If there is a significant number of disallowed combinations, consideration should be given to splitting the test design into separate stand-alone designed experiments.

☐ f. Review historical data

☐ g. Review known limitations to test

Date Completed: _____

☐ 9. For the response variables selected, estimate the following with 01B CTF and divisional analysts:

☐ a. Standard deviation (variability) of anticipated data for continuous variables.

NOTE: Preferably the standard deviation is estimated from historical or DT data. There are methods to estimate it based on the expected range of data, if required.

☐ b. Anticipated distribution of data (continuous normal, continuous but skewed such as χ^2, binomial (pass/fail), Poisson (small discrete integer values), etc.)

☐ c. Effect size and its operational relevance

NOTE: Effect size is related to the sensitivity of the test and can be thought of as the difference in performance that the warfighter will care about and that a statistical analysis needs to be able to detect in the data (if that difference is actually present).

This effect size may be the difference in performance between factor levels (e.g. high vs. low altitude bombing accuracy) or from a threshold (margin above or below where performance is critical).

Date Completed: _____

☐ 10. Meet with the divisional analysts, 01B CTF/DOE support to discuss and review inputs into the DOE calculations

Date Completed: _____

☐ a. Following this meeting, 01B will provide the following for each response variable:

i. Recommended type of statistical test/ analysis method for each response variable

1. DOE Analyses: analysis of variance, regression, response surface modeling, logistic regression, etc.
2. Non-DOE Analyses: one-sample t-test, binomial test of proportions, etc.

ii. Draft Hypothesis statements

1. Explain rationale for hypothesis test selected (e.g. one sided hypothesis test against a requirement) and an explanation of why they are acceptance or rejection based.

iii. Proposed run matrix

iv. Power/sample size calculations including confidence, power, sample size, standard deviation (if applicable), effect size, and any other amplifying notes and assumptions (to be incorporated as notes within the DOE run matrix table, Table C-1 Vignette to Subtasks to Conditions matrix).

☐ b. Additionally, 01B will calculate power associated with the highest priority factors, as applicable

Date Completed: _____

Date 01B DOE Received: _____

☐ 11. Draft the relevant statistical design paragraphs of the IEF for each response variable.

a. Explain/describe the response variable and why it is critical. Tell the reader what the threshold value is.

b. State the type of test that will be performed (acceptance/rejection) and explain why that was chosen.

c. Explain the test hypothesis, first in words, then in equations.

NOTE: ***Having already identified the response variable and the threshold value, use that language through the rest of this section, instead of referencing the null hypothesis or the alternative hypothesis, etc.***

d. Explain the operational relevance of the effect size and how it was chosen.

e. Include a discussion of the conditions/factors chosen and their operational relevance. Controlled, recordable and constant conditions should be identified as well as an explanation of any disallowed combinations

f. State the resulting initial design. This may include a matrix showing the factorial combinations that will be used. All assumptions should be addressed.

g. State the power, confidence, standard deviation (if applicable), effect size, and sample size associated with each response variable for the test design.

Date Completed: _____

☐ 12. Review the OTD Manual (408), schedule and conduct the DOE Working Group with 01B & Division A/B

☐ a. OTD documents action items and shares with attendees

Date Completed: _____

☐ b. Action items closed

Date Completed: _____

☐ c. Schedule E-IPR

Date Completed: _____

OTD IPR #2 Checklist

Purpose: This checklist should be used after an Executive IPR (E-IPR) has been conducted. It walks through the major steps needed to; develop vignettes, identify resource requirements and prepare for IPR #2. Products developed during these steps include the vignette to subtask to conditions matrix, the vignette to data requirements to test method matrix and the test event resource matrix.

Following the closure of any action items from the E-IPR, OTDs should have an approved task hierarchy for each COI. The task hierarchy should be linked to measures and conditions, with data requirements identified for each measure. These are required in order to proceed with the development of vignettes.

At the conclusion of IPR #2, OTDs should have completed the vignette matrices included in Appendix C. Test schedule, resources, limitations to test as well as modeling and simulation requirements will also be fully described. The majority of the IEF should be drafted and near readiness for routing for formal review.

Use of the IEF database is required for all IEFs. Most of the tables in the IEF document are produced directly from this database.

CAUTION: Using this checklist does not absolve the OTD of the responsibility for thought or the requirement to understand why they are performing these steps and how to apply the results. If you don't understand, ask.

NOTE: Vignettes are logical groupings of tasks/subtasks that support executing test events and data collection. While not in and of themselves a detailed test plan, they provide readers with a description of the type of test events OT will execute and the data gathered during each event.

☐ 1. Review the task hierarchy, conditions, measures and the linkages between them.

☐ 2. Obtain DT test plans if available and review them for specific test events that could support gathering of OT data.

Date Completed: _____

3. As an OTD begins to build vignettes there are several things to keep in mind:

☐ a. DT test plan & schedule (what functionality of the SUT will be available when, what DT's test events look like and how they intend to collect data against the MOEs/MOSs).

NOTE: Vignettes should be built with an eye towards reducing data collection requirements in IOT&E. If OT data collection can be conducted during a DT or IT event, build vignette(s) consisting of tasks that can be assessed prior to the OT phase. Review the task hierarchy and associated measures, and identify those tasks that can be assessed solely in DT/IT phases, those that will be assessed during all phases, and those that can only be assessed during an OT phase of test. Consider creating different vignettes for each.

☐ b. Location of test events

NOTE: The tasks that are able to be executed in a vignette developed to be conducted in a lab, systems integration lab (SIL), or using M&S may be different than the tasks an operator would execute in a real operational environment. Normally this requires creating separate vignettes for each.

☐ c. Resource availability

NOTE: OTDs should consider the availability of resources when building vignettes and may create different vignettes as a result. For example, an AW engagement has a clear Detect to Engage sequence of tasks, but due to missile availability, this could be broken into a vignette covering the Detect to ID tasks that could be run numerous times under a broad set of conditions, while a second vignette captures the Engage tasks associated with an actual missile shot under a narrower subset of conditions resulting from the DOE.

☐ d. Look for logical groupings of related tasks that can be combined into an executable test event.

NOTE: At the lowest level, a vignette could be constructed around one low level subtask. On the other end of the spectrum, one vignette may be created to describe an end-to-end test event that captures every task under a COI.

☐ 4. Using the IEF database, group related tasks that flow together into logical test events as a basis for each vignette.

☐ a. Review the task hierarchy and associated measures, and identify those tasks that can be assessed solely in DT/IT phases, those that will be assessed during all phases, and those that can only be assessed during an OT phase of test. Consider creating different vignettes for each.

☐ b. Meet with the 01B CTF to discuss vignette options. Building vignettes is a subjective process and can be done many different ways. Ask for and review examples from previous programs

5. Examples:

   a. Create a vignette capturing all the subtasks under one 1st level task, e.g. “Prepare/ Configure” (the first 1st level task for multiple COIs). This group of tasks may lend themselves to a mission planning test event for an aircraft SUT or a loading/prepare to embark vignette for a surface vessel.

   b. Create a vignette that captures all the tasks associated with multiple 1st level tasks, e.g. Search, Detect, Track, ID (common 1st level tasks for multiple COIs). This group of tasks may be logically conducted in one test event.

   c. Create a complete end-to-end vignette for use during IOT&E that captures every task needed to use the SUT for a particular warfare area/COI from start to finish.

Date Completed: _____

☐ 6. Title and number each vignette. Vignette numbering is based on the phase of test it will be executed in (IT or OT), the primary COI it relates to (#), and the number of vignettes (phase of test-COI-vignette #). Examples = IT 1-2, OT 2-1, etc.

NOTE: Vignettes may cross COIs – they are not limited to tasks described in one COI. Tasks for multiple COIs that are performed simultaneously or tasks that are common to multiple COIs may be combined into one test event. (e.g. maintenance tasks and data collection for maintenance metrics may be captured in multiple/all vignettes).

Date Completed: _____

☐ 7. If the SUT is net enabled it requires an IA COI. Ensure the appropriate IA vignettes are captured.

   a. At a minimum, for all systems that exchange information with another system, an Operational IA Vulnerability Evaluation (OIVE) vignette that looks at the system’s protect, detect, react and restore capabilities needs to be included.

   b. If the SUT has an IA COI, a penetration test vignette will also be required.

   c. Consult with 01C, for additional guidance.

NOTE: Penetration testing is conducted by NIOC and must be scheduled 1 year prior to testing.

Date Completed: _____

☐ 8. See the OTD Manual (405, Step 10), the Analyst’s Handbook (chap 8), and the IEF template for additional guidance.

☐ 9. Verify that the draft vignettes logically address and relate to the major test objectives for the SUT.

☐ 10. Ask the 01B CTF to verify that all tasks in the task hierarchy are linked to a vignette. This is a straightforward automated cross-check that can be performed. It won’t tell the OTD if the linkage makes sense, but it will identify any tasks/subtasks that have been overlooked and need to be mapped to a vignette.

Date Completed: _____

☐ 11. Review the draft Vignette-to-Subtask-to-Conditions Matrix for each vignette produced by the IEF database.

NOTE: This table identifies the subtasks captured by each vignette, the controlled and recorded conditions that apply to it, the run matrix and any DOE notes that may apply.

CAUTION: Depending on the number of controlled conditions, and the presence of response variables associated with the tasks covered by each vignette, the run matrix must be tailored and manually inserted into the table. Following the DWG, the 01B CTF should provide OTDs with this piece for each response variable. Seek their assistance in building these tables.

☐ a. Using the task hierarchy, review the conditions associated with each of the vignette tasks. Note that there may be some conditions associated with the tasks that did not apply to the response variable and weren't used for the DOE (if applicable), but still apply to the vignette. Identify these conditions.

Date Completed: _____

☐ 12. Add additional runs of the vignette to the run matrix to demonstrate the execution of the vignette under those conditional variations as necessary. Do this with 01B CTF support. Consideration may be given to creating separate vignettes for these runs (i.e. if they'll be executed during a different phase of test).

Date Completed: _____

☐ 13. Insert any applicable DOE notes for each vignette into the Vignette-to-Subtask-to-Conditions Matrix. These notes will be provided by the 01B CTF.

Date Completed: _____

☐ 14. Draft the relevant Test Execution sections of the IEF:

☐ a. With the SUT integrated master schedule, review the planned phases of test (DT/IT/OT) and draft the overall Operational Evaluation Approach section of the IEF. This section should lay out the major phases and provide a top-level description of how OT will participate in each.

☐ b. Draft the OT Vignette Strategy section of the IEF. Identify and briefly describe the vignettes that will be used by OT to assess performance of the tasks and collect data

☐ c. Draft the Schedule of Events section of the IEF. This should outline when the vignettes are expected to be executed (during which test period or test phase). Creation of a table is recommended.

Date Completed: _____

☐ 15. Produce the Vignette-to-Data Requirement-to-Test Method Matrix from the IEF dB and edit the Data Requirement block as needed for each vignette.

NOTE: Review and update the data requirements identified as part of the IPR1 checklist. Data requirements should provide the data element (Temp), the unit of measure (degrees F) and the source. Qualitative data requirements need to be defined with enough detail to support the creation of surveys (source) included in the test plan.

☐ a. Group data requirements under logical headings vice simply providing a list (i.e. group them by surveys, data sheets, significant tasks or capabilities covered by the vignette). See the IEF template for an example.

CAUTION: The export from the IEF dB will pre-populate the data requirement field with all the data requirements for the measures associated to that particular vignette, but it requires manual manipulation by the OTDs to logically group them (e.g. by surveys, DT only data, or other categories). Be sure not to delete any. The data requirements for every measure need to be identified here.

☐ b. Verify data requirements for each measure associated with that vignette are captured in this table.

☐ c. For IA vignettes, consult with 01C and review the default Vignette-to-Data Requirements-to-Test Method matrix available IA test plan template on the Y: drive. (Y:\01C\Best Practices\IA Best Practices\best practice IA Test Plan Template ….).

☐ d. Review each vignette and verify that the data required to address all attributes/measures has been identified.

NOTE: The association of data requirements to vignettes (test events) is a useful reference for OTDs when developing surveys, logs, and data sheets to support data collection.

Date Completed: _____

☐ 16. Populate the Test Method field in the Vignette to Data Requirements to Test Method matrix.

NOTE: This is best done outside of the excel table for formatting purposes, and then pasted back into the table.

☐ a. The test method field should provide a top-level description of how the OTD is going to execute the vignette, what will happen during the test and a summary of where/how the required data will be collected. It should also be traceable to the data requirements presented in the same table.

☐ b. Organize the test method narrative into "Pre-Test", "Test Execution" (further subdivided as necessary), and "Post Test" headings.

☐ c. For IA vignettes, reference the Vignette-to-Data Requirements-to-Test Method matrix available in the IA test plan template on the Y: drive. (01C—Best Practices—IA Best Practices)

Date Completed: _____

☐ 17. Review the draft Vignette-to-Data Requirement-to-Test Method matrix as a whole.

☐ a. Verify the subtasks the OTD selected actually apply to the vignette. Incorrectly associating a task to a vignette can bring along measures, conditions and data requirements that do not actually apply to that vignette.

☐ b. Verify the data requirements accurately describe the data element and the source of the data

- [ ] c. Verify the test method narrative describes not only the things OT personnel will do to execute the vignette, but also how the required data will be collected.
- [ ] d. Compare the Test Method write-up to the run matrix in the Vignette-to-Subtask-to-Conditions matrix to make sure they are consistent.

Date Completed: _____

- [ ] 18. Determine Resource Requirements for each vignette
    - a. For each vignette, identify the following
        - i. Test articles – number of full/partial systems, specific configuration of the test articles, etc
        - ii. Test sites and instrumentation - ranges, labs, unique instrumentation, length of time required
        - iii. Test support equipment
        - iv. Test targets and expendables - # and type
        - v. Operational test force support – this can be as little as an OTD & analyst, or as much as ships, subs, aircraft. The number of flight hours, days of at-sea time, etc should be captured here.
        - vi. Simulations, Models and Test Beds – any M&S requirements including labs, software models/applications, pre-faulted modules for M-Demos, etc

        - vii. Manpower & Personnel Training – type and number of personnel required with the associated training/expertise requirement
        - viii. Special Requirements – Any other non-instrumentation type of requirements (databases, special data processing capability, etc.)

NOTE: Fleet service requests must be submitted to 01C fleet schedulers NLT 9 months prior to test.

- [ ] b. Populate the IEF dB with the above resource requirements for each vignette

Date Completed: _____

- [ ] 19. Perform a sanity check of all the vignettes. Having the data requirements, test methods, run matrices, resources and phases of test the vignettes will be executed in, consider whether any of the vignettes should be broken up (or combined) further.

Date Completed: _____

- [ ] 20. Determine Resource Requirement by phase
    - [ ] a. Using the resource requirements for each vignette and the proposed schedule of vignettes in the OT Execution Strategy developed earlier, identify the test resource requirements for each test period (DT-B1, IT-B2,OT-B2, OT-C1, etc).

NOTE: This will serve as the basis for the identification of resource requirements and input for the TEMP. Enough granularity needs to be provided in the IEF to support the generation of OT funding requirements documented in the TEMP.

☐ b. Create the Test Event Resource Matrix in the IEF.

Date Completed: _____

☐ 21. Having reviewed the data requirements, test methods and resource requirements while building the vignettes, write the limitations section of the IEF

☐ a. Review the guidance contained in the IEF template

☐ b. Draft a paragraph for each severe, major or minor limitation

NOTE: The limitation wording is slightly different depending on the phase of test. Limitations for OAs relate to assessing risk, while those for IOT&E relate to resolving COIs.

☐ c. Ensure the write up addresses the impact of the test limitation and any steps taken to mitigate it.

Date Completed: _____

☐ 22. Write the Modeling and Simulation section of the IEF

☐ a. Review the M&S guidance contained in the IEF template

☐ b. Identify each model/simulation called for in the vignette resources and discuss them in their own paragraphs.

☐ c. Address the plan for verification, validation and accreditation. See 01C for all M&S assistance. Review COTF INST 5000.1B for additional guidance.

Date Completed: _____

☐ 23. Meet with the 01B CTF and review the IEF products generated.

CAUTION: Prior to generating the final vignette tables from the IEF dB, check each one, verifying that all measures have been set to test or no-test. This is a common pitfall—if through the iterative process, links of measures to tasks are modified, this may impact the vignettes if they're already built or drafted. For example, if a vignette is built and formatted and is considered complete, but an additional measure is linked to a task, that measure is also added to all vignettes that cover that task. Someone needs to set it to test or no test for each vignette or the measure and data requirements associated with it won't be displayed when the vignette is produced.

☐ 24. Share the vignette tables with external stakeholders (DOT&E, program T&E IPT, T&E WIPT)

Date Completed: _____

☐ 25. Schedule IPR2 with 01B and divisional A/B codes

- ☐ a. Build brief using IPR2 template (available under the 01C best practices\briefs folder).
- ☐ b. Provide the following as read-aheads prior to the review:
  - i. Vignette to Subtask to Conditions Matrix
  - ii. Vignette to Data Requirements to Test Method Matrix
  - iii. Test Event Resource Matrix
  - iv. DRAFT IEF (if you don't have it, don't go)

☐ 26. Conduct IPR2

Date Completed: _____

Attendees:

☐ 27. Following IPR 2, close out any action items, finalize the draft IEF and route for signature.

- a. Seek 01B assistance for generation of the final excel tables to be included with the routed IEF. (01B runs macros for the OTD that format the tables appropriately.)

Date Completed: _____

28. Reviewed by OTC and routed to 01B for comments

Date Completed: _____

29. Comments received from 01B _____

Date Completed: _____

30. Comments Incorporated & Routed to Division A/B codes

Date Completed: _____

31. Comments Incorporated & Routed to Editors

Date Completed: _____

32. Editors complete and routed to 01A

Date Completed: _____

33. Comments incorporated and routed to 00D

Date Completed: _____

34. Comments incorporated and routed to 00

Date Completed: _____

Change	Date Incorporated
Updated Fishbone Diagram on Page 1	15 June 12
Removed and updated all references to PINs— guidance has been incorporated into OTD manual	11 Dec 12
Clarified A/B codes should be invited to TP2	11 Dec 12
Added guidance for review and creation of Table 2.1 at IPR1	11 Dec 12
Additional administrative corrections (typos, clarified wording)	11 Dec 12
Added guidance regarding OTD review of JCIDS documents to the very first step	20 Feb 13